**EGI.eu**

EGI 1st User Virtualisation Workshop

|  |  |
| --- | --- |
| Document identifier | EGI-UVW1-Minutes-535-V2 |
| Document Link | https://documents.egi.eu/document/535 |
| Last Modified | 26/05/2011 |
| Version | 2 |
| Group Acronym | EGI.eu PDT |
| Group Name | EGI.eu Policy Development Team |
| Contact Person | Sy Holsinger, Sergio Andreozzi, Damir Marinovic, Peter Solagna |
| Document Type | Minutes |
| Document Status | Draft |
| Approved by | Steven Newhouse |
| Approved Date | 27/05/2011 |

|  |
| --- |
| Abstract  Detailed minutes of presentations, break-out session discussions, afternoon summaries and final points of actions from the 1st User Virtualisation Workshop held 12-13 May 2011 in Amsterdam, NL. |

Copyright notice

Copyright © EGI.eu. This work is licensed under the Creative Commons Attribution-NonCommercial-NoDerivs 3.0 Unported License. To view a copy of this license, visit http://creativecommons.org/licenses/by-nc/3.0/ or send a letter to Creative Commons, 171 Second Street, Suite 300, San Francisco, California, 94105, USA.

The work must be attributed by attaching the following reference to the copied elements: “Copyright © EGI.eu (www.egi.eu). Using this document in a way and/or for purposes not foreseen in the license, requires the prior written permission of the copyright holders. The information contained in this document represents the views of the copyright holders as of the date such views are published.

Application area

This document is a formal EGI.eu report applicable to all participants and associate participants, beneficiaries and Joint Research Unit members, as well as its collaborating projects.

ORGANISATION SUMMARY

To support science and innovation, a lasting operational model for e-Infrastructure is needed − both for coordinating the infrastructure and for delivering integrated services that cross national borders. The objective of EGI.eu (a foundation established under Dutch law) is to create and maintain a pan-European Grid Infrastructure in collaboration with National Grid Initiatives (NGIs) in order to guarantee the long-term availability of a generic e-infrastructure for all European research communities and their international collaborators.

In its role of coordinating grid activities between European NGIs, EGI.eu will:

* Operate a secure integrated production grid infrastructure that seamlessly federates resources from providers around Europe
* Coordinate the support of the research communities using the European infrastructure coordinated by EGI.eu
* Work with software providers within Europe and worldwide to provide high-quality innovative software solutions that deliver the capability required by our user communities
* Ensure the development of EGI.eu through the coordination and participation in collaborative research projects that bring innovation to European Distributed Computing Infrastructures (DCIs)

The EGI.eu is supporting ‘grids’ of high-performance computing (HPC) and high-throughput computing (HTC) resources. EGI.eu will also be ideally placed to integrate new Distributed Computing Infrastructures (DCIs) such as clouds, supercomputing networks and desktop grids, to benefit the user communities within the European Research Area.

EGI will collect user requirements and provide support for the current and emerging user communities. Support will also be given to the current heavy users of the infrastructure, such as high energy physics, computational chemistry and life sciences, as they move their critical services and tools from a centralised support model to one driven by their own individual communities.

The EGI community is a federation of independent national and community resource providers, whose resources support specific research communities and international collaborators both within Europe and worldwide. EGI.eu, coordinator of EGI, brings together partner institutions established within the community to provide a set of essential human and technical services that enable secure integrated access to distributed resources on behalf of the community.

The production infrastructure supports Virtual Research Communities − structured international user communities − that are grouped into specific research domains. VRCs are formally represented within EGI at both a technical and strategic level.

TABLE OF CONTENTS

1 Context 4

1.1 Participants 4

1.2 Aim/Goal 4

2 Day 1 5

2.1 Welcome and Vision 5

2.2 Validating, Identifying and Endorsing the Use Cases 7

2.3 Break-out Sessions 10

2.3.1 VM Management Session 10

2.3.2 Information Service Session 12

2.3.3 Accounting Session 15

2.3.4 Monitoring Session 17

3 Day 2 19

3.1 Day 1 Summaries 19

3.1.1 Monitoring 19

3.1.2 Accounting 20

3.1.3 Virtual Machine Management 22

3.1.4 Information Services 23

3.2 Key points summary 25

3.3 Technical discussion: Implementation roadmap 27

3.4 Concluding Remarks 30

# Context

This document contains the raw notes (edited for readability) with no analysis from the workshop to provide a permanent record of the discussions. This document, alongside the presentations and discussion that took place at the workshop, will be used as a basis for building EGI’s Cloud related strategy.

## Participants

Over 70 participants were in attendance.

Participation list: <https://www.egi.eu/indico/confRegistrantsDisplay.py/list?confId=415>

## Aim/Goal

The dedicated workshop on user virtualisation was to bring together three critical groups within the European production infrastructure – end user and resource and technology providers.

The workshop was to address specific questions through a series of user specific presentations and topical breakouts covering monitoring, accounting, VM management and information system. The workshop was to serve as a follow-up to a recently published report on the integration of clouds and virtualisation into EGI (http://go.egi.eu/258) and a technical roadmap defining the EGI cloud profile.

Obtaining feedback from the community was essential for understanding if and how EGI should move towards providing an Infrastructure as a Service (IaaS) model to support data intensive research communities and collate a set of critical observations that could be developed into a roadmap that will meet the growing need for virtualised resources from the European research communities.

# Day 1

## Welcome and Vision

Steven Newhouse (SN), Workshop Chair and EGI.eu Director gave an introduction presentation providing an opening welcome address, outlined the EGI vision, and objectives of the workshop.

A few questions motivating the meeting:

* How to increase the diversity of user communities?
* Scale out to different environments?
* Virtualisation internal to the data centre?
* Virtualisation just for providing different batch execution environments?
* Expose and make virtualised resources accessible as with other resources?
  + Yes, but how?

Where to start?

* Federating virtualised resources as with the current EGI model

*Interjected question/comment*

* Jeff Templon (JT): Slide 6: “The Virtualised Future” diagram: The red part in EGEE included services like WMS, BDII, etc. With this new scenario, the red part becomes closer to bare metal. People in the yellow box need to manage also the deployment of services that were managed in the red box level before. Do we have this people and who wants to pay for?
* SN: Yes, within the NGIs. What is the use case of accessing the lower layer? We have the 6 basic scenarios from Cloud architecture document; having a way of plugging this in.
* JT: We need to decide what services go where, define the line.
* David Wallom (DW): Communities like ELIXIR have built experience and software for years. They want to leverage the experience and products.
* Owen Singe (OS): Confused about trust here, is about the end-user?
  + SN: Different levels, starting the virtual machine (VM) is one level; accessing the service within the VM is another level; many different trust relationships are building up.
  + OS: I don’t see how expert users can give certain level of trust to non-expert users.
  + SN: Trust model is provided through the gateways. Can they handle or manage these new trust models? This issue to be addressed later in the meeting.

*Presentation continues*

* No big bang migration - gradual change transparent to the end user.
* Clearly identify the role of the expert.
* Increase flexibility of infrastructure.
* Leverage expertise throughout the community.

6 Key Usage Scenarios

* Running a pre-defined VM image.
* Running my VM image (with my data).
  + Certification process needed to: upload the VM image/upload data into the VM/turn it on and use it.
* Deciding which virtualised resource to use.
* Accounting across resource providers.
* Reliability/availability of the resource.
* State change notification from the VM manager.

Questions to be answered in breakouts:

* Are there standards we should be using?
* Are there best practices we should adopt?
* What is the maturity of the software?
* What is the availability of the software?
* What are the priorities?
* Where are the gaps, issues & concerns?
* What work is needed to remove these?

Feedback:

* Need to decide if these are the right scenarios to get started, what are we missing, what can be added. (But not attempting to come up with an exhaustive list).
* Need to hear back from Technology Providers and if they can address concerns.
* Resource providers for engagement in a testbed of virtualised resources and how does integration in to EGI work?
* Response from end users (or reps) and interest and ability to collaborate.

*End of Presentation - Floor opened*

Erwin Laure (EL): In the area of accounting, so far EGI is kind of free of charge; if you go to commercial cloud, somebody needs to pay the bill; have you modelled how to charge/pay?

* SN: JRA1 will work in this area to derive billing models. I would suggest to not address in this workshop, but in a later one planned specifically on this.
* DW: It is not actually free; somebody has to pay.
* OS: Address from different perspectives; accounting can be used to limit the usage of the resources; how do we keep fair share access to resources as we tried to do for years? The site on one side wants to maximise the utilisation, while the other side wants to guarantee fair share.
* JT: In the VM scenario, how do we make users release the VMs and not keep them for months?
* Ian Bird (IB): Schedulers are used even for virtualised resources; fair share can be included there easily.
* DW: You can keep a VM for 3 months; we will just charge users for VM usage.
* SN: The difference between the first two scenarios is how much control you give to users to upload their VM.

Response to the posed questions to be answered for the breakouts:

* JT: Are we limited to these questions? From CMS we got different requests:
  + As a site provider, can we monitor what is happening within the VM?
* DW: Should be separated in the monitoring of the status of a VM, and the status (accounting) of the service running on it.

General comments:

* Neil Geddes (NG): Asked a question regarding datasets.
* DW: There are services, which fit well in VM, others don’t.
* JT: For the VM session, you may need to be able to specify the connectivity; if you have a Monte Carlo job, you do not need hi-speed connectivity; but if you application that does need high performance access to data, you want to know which VM and host environment can provide it; Requirement for information service and VM management.
* EUDAT project started; many organisations participating.

## Validating, Identifying and Endorsing the Use Cases

Participating projects and user communities validated the EGI Cloud Architecture Use Cases by presenting overlap and disparity of their own use cases:

* Which project/community use cases fit well with the EGI Cloud Architecture Use Cases?
* Which project/community use cases do not fit in?
* This is an open session where projects and communities can align their planned use of clouds to the 6 identified scenarios and where their planned use goes beyond these 6 scenarios.

Presentations included:

* EGI Virtualisation Six Scenarios
* StratusLab use cases
* CSC-ELIXIR use cases
* WLCG Trusted Virtual Images
* WLCG Use Cases

**CSC ELIXIR Scenario** - Michel Drescher (MD) presented on their behalf

For content see .ppt available on workshop agenda page.

*Questions from the floor*

* Luděk Matyska (LM): How is this model dissimilar to the grid model we are using?
  + SN: It is another approach: having a local cluster, expanded (scaled out) into an external virtual cluster.
  + MD: The end-user does not see any difference; it is transparent.
* OS: Image we have popular sites, they are usually fully subscribed; how do we deal with that?
  + SN: That is why we have monitoring and registries.

**WLCG Views on Clouds** - Presented by Ian Bird:

* ATLAS: The most advanced in this area. They would like to evaluate, design a model to interact with their software.
  + It is more efficient to give a multi-core machine to a user and let him optimise it instead of hiding it.
* Questions:
  + Who build the VMs? Who instantiate the VMs?
* CMS: They are less interested. They see VMs as a site business; As for monitoring, they have it at the job level, so no difference.
  + Many HEP groups made experiments running on Amazon with different results.
  + They would be happy with commercial clouds if cost of ok.
* LHCb: Aim to replace Dirac pilot with a customised CernVM, using CernVM certified images.
  + CVMDirac could run on Amazon.
  + LXCloud or similar institutional cloud - as alternative to batch system.
  + Opportunistic usage - e.g. BOINC can also be used with multi-core usage; some development needed.
* Site use cases:
  + Experiments with OpenNebula and Platform ISF.
  + WLCG - How far can standard cloud interfaces replace grid job management?
* EC2 is the current de facto standard. We should think about it before generating new standards.

*Questions from the floor*

* SN: It seems that you see two different approaches/interfaces when instantiating on Amazon vs. private cloud. Any discussion on this.
* IB: Not in depth. They will be happy to do “a la amazon” also on private cloud.
* IB: Policy question: if we go to commercial cloud (e.g. Amazon) can we use research network?
* SN: Amazon gives a service and the users must accept it as it is. The cloud/virtual services provided by the grid community could be adapted to experiments needs.

**WLCG Trusted Virtual Images** - Presented by Michel Jouvin (MJ):

* Trusting, certification of virtual images.
* Revocation looks to be harder than endorsement.
  + For example, the StratusLab marketplace has an expiration process for images; they are not endorsed forever.
* Scale issue in updating images; to patch, re-endorse.
  + Every small update on the VM operating systems turns to be a new image to endorse/distribute.
    - Is up to the sites to accept systematically all the images from a certain user (VO manager).
  + If VM images include also the exp. software we would have a new VM image every day. That is not possible. But the experiment software could be accessed through a network file system mounted on the VM.
  + Site admins can have different policies on accepting images, for one endorser site could accept all the VM, for another endorser choose VM by VM.
  + How to remove VM images which cannot be trusted anymore?

**StratusLab Scenarios** - Presented by Vangelis Floros (VF):

* Matching EGI cloud integration profile scenarios with StratusLab’s use cases
  + All EGI scenarios can be mapped in one way or another, either matching or extending each scenario.
* New Scenario: Credit-based resource utilisation (priority: need later)
* New scenario: Transparent access to federation resources (priority: need later)

For further details see .ppt available on workshop agenda page.

*Questions from the floor*

* JT: Sounds like use of passwords, I thought we were moving away from this.
  + VF: We only have one site, so we are currently investigating this for when we have a federation of sites, using for example VOMS for authentication is an option.
* DW: What do you mean for “user” in this case?
  + SN: People instantiating a VM.

**ISGC Paper Model: Grid and Cloud integration scenarios** - Presented by SN, slides from Alberto Di Meglio)

* Scenario 1: Cloud resources behind grid services
* Scenario 2: Shared grid and cloud access
* Scenario 3: Shared access with dynamic brokering
* Scenario 4: Dynamic grid services

For further details see .ppt available on workshop agenda page entitled: EGIUVW1.pptx

*Questions from the floor*

* JT: Scenario 3 - We are ignoring the data, this scenario takes starts to take care of this.
* Models 2, 3, & 4 mix different kind of trust: trust to deploy VMs, trust to access virtual resources. Scenario 1 has the same level of trust.

## Break-out Sessions

### VM Management Session

Chair: Alexander Papaspyrou (AP)

Questions:

* Are there standards we should be using?
* Are there best practices we should adopt?
* What software is available at the market?
* What is the maturity of the software?
* What is the availability of the software?
* What are the priorities?
* What are the gaps, issues, and concerns?
* What work is needed to remove these?

Discussion items:

* Specify bandwidth for a VM for data-intensive applications.
* Expire and revocation of images (probably out of scope).
* Trust, endorsement, approval policy and model (certainly out of scope).
* Authentication mechanisms (preferably X509v3 or VOMS proxy).
* Bulk operations.

*Discussion*

* Morris Riedel (MR): Grid appliances and other appliances (community) - Differences, relation to one another.
* Grid appliances are the traditional middleware stack packages as VM.
* Community provided VMs.
* From VM management?
  + In general no; maybe on regularity of updates.

What is VM management?

* At what level?
  + Someone needs to address policy issues, (e.g. you can run a VM for a specific time) - not to be discussed here, but a recommendation.

Requirements:

* HIGH: Parameters to instantiate a single VM.
  + Bandwidth for a VM (for data-intensive applications), network type memory, running time, disk space, IP address, # cores, CPU, firewall rules, running time.
* LOW: SLA - exclusive usage of physical machine, requested uptime.
* MEDIUM: Instantiate *n* VMs as cluster.
  + Shared disk space, connectivity.
  + Dependencies between VMs:
    - Deploying one VM implies deploying other VMs.
    - Identified by a tag.
    - Provisioning VMs should be focused on the infrastructure level, anything further down is the problem for the communities.
  + Leads to two levels of deployment scenarios.
* LOW: Do we need advanced management (advanced reservation?)?
  + HIGH: In a testing environment (maybe nothing now, more for the future).
  + VO and site level needs to be able to schedule.
    - Load balancing and fair share.
  + Does the interface need scheduling functionalities?
    - Need it internally, but via the interface?
  + NB: If I give you a VM, we must be able to ask, how long will you have it?
    - From request to instantiation - for single and bulk submission.
    - I need 500, can I even get 500, tell me, cause if not, I will go to another site where I can - a capability we don’t have at the moment - good time to introduce this.
      * In a federated scenario, makes sense.
* MED: Expiry and revocation of images.
  + Revocation can be done now by policy decisions at the site - covered by EGI scenario 6 trusting images. Is similar to trusting physical machine. Restart policies will be different throughout the sites.
    - Can only start golden images, then limit the run time, and restart from only the original trusted image.
    - Sites are more trusted to let user restart VMs from saved snapshots of images.
  + Need to have a notification of bad images.
    - Balance risk and usability - we have mechanisms in place.
    - So we need formal information to differentiate between the two.
* HIGH: State view from the VM management interface of what is running on the system.
  + A user wants to know which VMs are running on a certain systems.
  + A sys admin wants to know who is running what.
* MED: A user can request a snapshot of its running VM. The API if enabled by the site.
* HIGH: Authentication mechanisms:
  + Preferably x.509v3 or VOMS proxy)
  + Evaluate SAML/Shibboleth, EDUGain...
  + 2 level of authentication (out of scope of this meeting).
    - Super-users starting and stopping VMs at a site.
    - Users of services running in a VM.
  + The decision of which hypervisors should be supported by sites is a policy decision.
  + VM management is agnostic to image format (supported hypervisors), but expose which hypervisors are supported.

Standards:

* VM Management:
  + OCCI for VM management (create, start, stop).
  + OVF for VM format.
    - OCCI and OVF need integration.
  + Moving VMs between sites
    - CDMI can be a possible solution.
  + Connecting OCCI/OVF/CDMI.
* EC2? Maybe higher level.
  + At the moment, it is the API on the resource provider side.
  + Messaging (out of scope).
  + Cloud from DMTF, similar goal of OCCI.
* Maturity?
  + TCloud similar to OCCI.
    - StratusLab has one implementation.
  + CDMI: reference implementation + SARA.
  + Move rest till tomorrow for other implementations.
* Networks.

State of VM management implementations:

* The requirements should be on the interfaces.
* What VM management solutions can we consider?
  + OpenNebula, OpenStack, Nimbus, Eucalyptus, batch schedulers (in the area of user jobs as VM), Platform ISF, VMware.
* What kind of VMs are we providing?
  + In batch system mode.
  + On-demand.
  + Differentiate between batch/HTP VMs and on-demand VMs
  + JT: They need to be there, cause that is what the new users will have had on their computers before.
* What is the difference between starting a process and starting a VM?
  + Should be faster using virtualisation.
  + RedHat provides features to tune process resource consumption.
    - RedHat has just announced its cloud architecture.
  + Processes can be check pointed, migrated, …

List general requirements now, then prioritise later.

### Information Service Session

Chair: Morris Riedel (MR)

Roles:

* Information System (model).
* Metadata for describing and using VM applications.
* Dynamic registries for dynamic services discovery.
* Technology (sustainable) and standards (maturity).
* Existing infrastructure integration (e.g. end-users).
* End-users to select appropriate services (and VMs).
* End-user information.
* Protected information - is there some need for it?

Potential out of scope or extras:

* Interrationships with dynamic configuration and other breakout sessions.

Scope:

* Services, VM images,...?
* We want to discover info about:
  + Available VM appliances.
  + Scheduling related parameters useful to perform decisions.
    - Hypervisor type, Virtualisation technology.
  + CPU/GPU, # cores, … (note from VM management: bandwidth, network type, memory, disk space, IP address, # cores, CPU, firewall rules, running time).
  + Authorisation info (Can I instantiate a VM through a certain endpoint?).
  + Pricing.
    - Business models and pricing is out of scope and to be moved to a different workshop. Now we must concentrate on getting the right information, making sure we can account for everything, whether they are virtualised resources or actual resources, is the first priority.
* To investigate overlapping of scheduling related info with monitoring requirements.
* Monitoring: Who is starting/running VMs?

Consumers of the Information:

* End-users.
* Brokers/meta-schedulers (e.g. WMS).
* Other services.

Standards:

* Information model:
  + Leverage glue 2.0:
    - Endpoint to discover VM management endpoints.
    - Execution environment to discover types of instances, where to deploy VM images.
  + OCCI defines key/value pairs.
  + OVF covers specification of VMs, network, and storage.
  + DMTF:
    - Virtual system profile.
    - Virtual system virtualisation profile.
* Information discovery interface:
  + OCCI has a discovery interface + resource description based on key/value pairs.
    - Some mandatory (e.g. related to # cores, CPU speed).
  + CDMI has a discovery service for data.
    - Can express capabilities of storage system (e.g. type of FS).

Best Practices:

* From cloud
  + StratusLab will address during 2nd year (federation of cloud systems).
  + Based on OpenNebula (related to monitoring):
    - Info about the physical system (what is available/used).
    - Ganglia for status.
* From grid
  + For info model, evaluate type of info captured in GLUE 2.0.
  + Difficult to upgrade information models due to dependencies on usage/sensors
    - Re-use experience and try to get it right.
    - Importance for info service, quality of data more than completeness.

Guidelines:

* Evaluate how to simplify the information needed at finer grain level.
  + They have been modelled in Grid, but somewhat difficult to capture in a meaningful way.
* “Freshness” of information; should we have a policy in the profile like “info should be old less than 5’”?

Software:

* How to deliver the information?
  + Messaging: ActiveMQ.
* How to aggregate, store, query info?
  + BDII, OpenLDAP, …
* Evaluate approach of repository as DB vs. in-memory.
* P2P vs. centralised vs. hierarchical.

Priorities for the Information Model

* What capabilities need to be represented?
  + Compute:
    - VM Appliances or as part of VM registries.
  + Type of VMs you can use.
  + Storage.
* Network.

Gaps:

* Isolation between users about usage of the infrastructure.
* Amount of resources available to a particular user.

Work needed to remove the gaps:

* About isolation of info, not high-priority.
* From the technical viewpoint, moving towards virtualisation does not show big problems.
  + Issue with revising the info model, the process can take time.
* Analyse overlaps and differences between grid and cloud standards related to info model/discovery.
  + Maybe leverage common meetings of SDOs.
  + Next week DMTF, OGF, SNIA will meet to discuss integration of OCCI, CDMI, and OVF.
* Minimum profile.

### Accounting Session

Chair: John Gordon (JG)

* Accounting does not look different than OGF accounting record.
* Operation of concern when somebody sends it to a VM and it is charged.
  + Double billing issue.
* As a recourse provider you don’t see internal accounting. You would want to know what use of the VM was. Double accounting does not account, double billing does. VM use exists.
* Quality of service issue for virtualisation.
* Commercial guys use different templates for different kind of things.
  + You pay for the limit. Power usage is a difference.
* If you overprovision cost will go up.
* If my SLA said I will provide you with the box providing.
* You need to be predicable. We can enforce that we cannot use more memory.
* We are accounting for service providers not for users.
* Capacity management.
  + You sell as much as you can at risk and you are hoping full capacity will not be used.
* Accounting for work done in a VM can be done by traditional methods.
  + As long as you make sure any logging you need is persistent beyond the end of the VM.
  + I don’t think we need to spend much, if any, time of this today.
* Accounting for the use of a VM instance is similar to that of a job, in that the resource you are accounting has a defined lifetime like a job and most of the resources you would account on like CPU, memory, i/o, networking, were foreseen in the OGF UR even though no-one has actually implemented them all yet (AFAIK).
* Extend PGF UR?
  + Both the above need to be done but one risks double billing.
  + e.g. both the VM owner and end user get charged.

Software:

* EGI has infrastructure in place to collect accounting records.
* Hypervisors keep logs.
* Amazon etc. produce bills.
* Other projects?
  + FleSSR
* When you got accounting you need to apply function and that is difficult. Any other software things?

Features and requirements:

* Accounting for VM.
* Are a VM accountable like a job?
* For VM (e.g. services) running for months you don’t want to wait for the shut down of the machine (as for grid jobs) to bill for it. You need intermediate bills (e.g. weekly) for that VM usage.
* If you account for a VM you account for all the time (with all the cores) the machine runs, either if it is idle or not.
* Accounting for storage.
* Accounting for a service more than the usage of the disk space. The persistence storage is reserved also after the VM is shut down.
* Currently there is no usage record for storage; as soon as it is available it could be used for both virtual and physical storage usage.
* Important assumption what is actually operation, can I can create separate volume and attach the new virtual instance with status and how big it is. When you attach, when you detached. User can use transitional storage but less amount of persistent storage. Recording a snapshot and added to usage record. We should look more for usage record of service. Storage accounting is about accounting storage not some bit. Persistence storage is the same that it was. Accounting service based accounting.
* These models are resulting in free services. If we have any success, our resources will be occupied. Than queues and fair shares. It is more about monitoring.
* Why to make distinction short and long running jobs?
* 5 different templates two different priority.
* How do you collect info?
  + We do it from the outside. Reservation is stored in database.
  + We don’t have storage usage record, but when you have one. Downside scalability issue. We overprovisioned information that is here.

Standards:

* Usage record is the standard used in many place.
* Intention is not create one big standard; it is to work on next generation of usage records.
* Push through standard based accounting and OGF fits.
* Basis of something on OGF usage record and we should push on it.

Use:

* Account the actual CPU power available for the VM. That is what the user gets.
* The provisioning of an instance defines the major part of the accounting, not the actual usage.

Gaps

* What do we need to do?
  + Network accounting.
  + Storage accounting.
  + Data usage accounting VM usage record.

Issues:

* Identity management.
* Billing.
* Normalisation.
* Fairness and capacity (out of scope).
* Agreements and policies.

### Monitoring Session

Chair: Jeff Templon (JG)

* Dennis van Dok (DD): There is already documentation to see what you need to monitor to consider your site secure.
* We should never look inside of VM? Most of the people agree that you can’t. Somebody should look at the machines.
* IB: You should be reasonable looking, not looking actively for it, when there is certain event, forensic issue. Pre-emptive looking is not possible or desirable.
  + Is it legal to look inside of VMs?
    - It all depends of your ToC
  + JT: Not a monitoring issue, more security issue but it should be recorded.
  + DW: Apply Amazon approach, who owns VM that one is responsible. No use case to look inside of machine.
* IB: If sites move infrastructure to be virtualised they should be no obstruction to obtain monitoring they now obtain Infrastructures. Dashboards should move to messaging service.
* JT: We need to ask users for monitoring they needs. Not to guess, since they are not here.
* Gergely Sipos (GS): The topic belongs to accounting. What is the difference between monitoring and accounting?
* System health, can users instantiate new virtual machines?

Issue is how to define scope of monitoring, lot of difficulties hot to define monitoring.

* JT: Monitoring is before VM starts. No use case to work inside, ill defined what the scope of monitoring is.

What do we need to monitor?

* Some of the things that need to be monitored are more specific to rest of us.
* Network activities need to be monitored. However, it is not quite clear what are networking activities.
* IB: Network monitoring is more important than before because there is less trust for what is going in VM.

Status of VM management layer:

* What does Amazon provides to users:
  + CPU utilisation.
  + Available disk reads.
  + Available disk writes.
  + Max and rate network in and out.
* DD: Reason we do monitoring is to avoid having a significant number of failures.
* JT: Current model we actually submit the job.
* Run a VM creation test. We need analogue of CE test.
* Need to see performance monitoring.
* It should be user-centric.
* Security?
* Discover what monitoring is in VM management layer and what is already there?
  + OS: How much it would cost me. That is VM management.
* Do we want community probes directed at detecting contention?

Answers to the needed questions at infrastructure level:

* Standard: De facto Nagios.
* Best practices.
* Maturity level of the software.
* Availability of the software - licences and platforms.
* Priorities (e.g. ignore network in/out separation).
* Gaps, issues and concerns.
* SN: Biggest gap is how to provide provision the life cycle. It is needed. We should build solid foundation on this workshop and other people can think what kind of high-level services need to be created.
* IB: Careful in defining the model of this, no user community to say what they need. They look for cloud infrastructure or grid?
* DW: Presented CLARIN diagram they contain persistent and scalable services.

Issues outside the scope of this session:

* There should be a unique handle that needs to verify whatever you started it is over.
* We assume that VM management layer is sufficiently reached to determine whether it is ok to shut down machines.
* There is a large class of things to be monitored for scheduling purposes.

# Day 2

## Day 1 Summaries

### Monitoring

What’s left?

* Determine whether the service is functioning correctly.
* Basic info on activities of VM:
  + Amazon provides:
    - Average and peak net in/out, CPU utilisation, and disk reads/writes.
* Understand whether there are contention problems with farm.
  + May be handled by VM management layer.

Open questions:

* What monitoring of VMs is provided by various VM management systems?
  + Is this enough?
* If not, do we want to develop community proves (ganglia?) aimed at detecting the metrics not provided directly by the VM management layer.

What to do:

* Service level tests of VM gateway service.
* External tests “submit a VM and see what it does”; akin to current job submit tests.
* Nagios framework already exists; use this.

Answers to the ended questions at infrastructure level:

* Standards:
  + De facto standard: Nagios.
* Best practices: Nagios.
* Maturity level of the software: OK.
* Availability of the software: OK.
* Priorities:
  + Develop the two plug-ins.
* Inform ourselves first on best practices.
* If nothing new is needed, develop nothing.
* To 0th order should need nothing new since a VM is just an application.

Other:

* Monitoring is just to know if the system is running correctly.
  + Rest of the monitoring is from the user perspective.
  + Issues identified but out of scope.
* Users want to know which VMs they have started and what is their status.
* We assume that the VM management layer is rich enough to have internal monitoring where necessary (e.g. can I shut down host z in ten minutes?).
* Large class of things to be measured for scheduling purposes. When will my VM run, how many VMs can I run, etc.?
* Where is the data closest to my job?
* DW: From the infrastructure point of view, a machine is successfully started when it has network connection and requested storage attached.

### Accounting

Scope:

* Accounting for work done by users inside a VM should be done by traditional methods. (out of scope)
  + e.g. an elastic WN will report its accounting back through a real or virtual CE.
  + Just make sure any logging you need is persistent beyond the life of the VM.
* Cloud Accounting should be for the use of a VM instance by the person (?) who instantiates it? (in scope)
  + If they wish to carry out more granular accounting within the VM for recharging that is their prerogative, but is out of scope.
* Storage Accounting:
  + Elastic Storage in VM. (in scope)
  + Persistent Storage (out of scope). Cloud and Grid Storage should cut the same accounting record.
  + Residual issue of adding storage accounting capability to purely cloud-based persistent storage.
* Network Accounting:
  + Some network use can be accounted to the VM, but things like virtual switches don’t fit there.

Accounting for the VM:

* The life cycle of a VM has many similarities with that of a job.
  + It has a defined lifetime and most of the resources you would account for (e.g. CPU, memory, i/o, networking). Was foreseen in the OGF UR even though no one has actually implemented them all yet (AFAIK).
* Majority view that one should charge for the footprint of the VM times lifetime.
  + e.g. # cores\*(endtime-starttime).
  + NOT actual CPU used by VM.
  + Can we have big enough clouds with sufficient over provision that an instantiated but inactive VM can attract zero or very low charges?

Software:

* EGI has APEL infrastructure in place to collect accounting records.
  + Extensible to include new record types like VM.
* Hypervisors keep logs; OpenNebula has tools to gather these.
* FleSSR has defined a schema of VM characteristics, which it has harvested from Eucalyptus.

Standards:

* Start from the OGF current standard: Usage Record.
  + Extend the current standards, if required, to account virtual resources.
  + SIENA Roadmap is to extend current standards (like OGF) for VM accounting.
* Standards for network accounting?
  + Account for reservation of resources and for actual usage.

Use:

* FleSSR has collected accounting records.
* No evidence of real work from other cloud projects although they all show willingness.
* Examples exist of local clouds accounting their use in various ways.

Gaps:

* VM usage record.
  + Kick off with OGF in existing UR-WG.
* Storage accounting.
  + Check whether there are virtual storage types to be instrumented.
* Network accounting (big gap).
* Data usage accounting.
  + # of I/O operations (who’s billed?).
  + Probably out of scope for cloud. Same solution as traditional data centres – no standard.

Issues:

* Standard hooks in Hypervisors to provide the information.
* Identity management:
  + Current infrastructures don’t use the same ones as grid.
  + Sites need to identify who endorses and instantiates images.
* Normalisation:
  + The power of a virtual CPU is not so well defined.
* Billing:
  + How long have you had it?
* Fairness and capacity (out of scope):
  + How is over commitment and pre-emption treated for accounting?
* Agreements and policies:
  + Framework for recharging agreements.
  + Policies for exchange of data, pricing definition.

Comments:

* Billing is not simple, should discuss lot more about fulfilment
  + To be discussed in greater detail in a later, dedicated workshop.

### Virtual Machine Management

Standards and best practices:

* Standard Marked well-covered:
  + De facto:
    - EC2.
    - Azure.
  + Open standards
* VM management developing (still developing) they are young, not much experience on them.
  + OCCI family and TCloud.
  + OVF for VM description.
  + CDMI companion data provisioning.
  + Network still unclear.
* Harmonisation required!

Action: Evaluate maturity for proposed products.

Software on the market:

* + Diverse market:
    - Commercial, academic, open source.
    - Classic VM hypervisors.
    - Cloud offerings.
    - Traditional batch system schedulers.
  + Should consider supported API standards to achieve a harmonised environment

Requirements and priorities:

* Deployment:
  + HIGH: Parameters to instantiate a single VM.
    - Including the running time decided in advance.
  + HIGH: API should expose supported hypervisors.
  + MED: Mechanisms for deploying VM landscape.
  + LOW: Specification of QoS (via SLA).

Management:

* HIGH: Bulk operations.
  + Not only for creation/shut down of machines, but for more operations.
* HIGH: State view.
* MED: Expiry and revocation of images.
* MED: Snapshot taking.

Security:

* HIGH: Support traditional (e.g. x.509, VOMS), but consider other (e.g. SAML, Shibboleth,...).
* MED: Provider should be able to understand, which running VM is based which image (for revocation).

Capacity planning:

* Scheduling capabilities in the interface.
* How long will it take from request to instantiation?
  + Higher priority in a production scenario.

These are concerns from the communities, people need these features, somehow, and interfaces/API should address these requirements.

Gaps issues concerns:

* Need to understand whether different types of appliances require differences in management.
* Networking seems to be underrepresented in the standards area.
  + What is missing is to get description of what’s guaranteed, descriptions for firewalls.
* Trust, endorsement, approval policy and model not entirely clear.

Questions from BigGrid:

* OVF already supports all kind of hypervisors. All the tools seem to be available; they just need to be integrated.
  + The tools are there, we need to check if all the needed fields are in place.
* Network working group in OGF.
  + As for the previous one, it has not been checked how it fits to the cloud requirements.
* OVF+ is the new version.

IB: Exclusive resources

* What makes more sense is to go to a provider that maps the virtual resources into the bare metal machines

### Information Services

Scope:

* Information about services in line within a VM instance.
  + Community specific (out of scope).
* Perspective:
  + Information about VM instances itself.
    - Infrastructure (in scope).
    - Information for scheduling (overlaps with monitoring).

Common understanding:

* Hypervisor type.
* Need careful investigation to understand overlaps with monitoring.
* Assuming there will be a submission of VMs to sites possible.
* Scheduling related parameters useful.
  + Different capabilities often related to compute ones.
  + CPU/GPU, cores, hardware architectures.
  + Machine speed, bandwidth, disk space, firewall, IP address,...
* End user information?
  + More scientific end-user input required.
    - Not much interest in that.
* All done in grid seems to be useful in cloud.

Other related work and topics:

* Available VM appliances -> image repository
  + Info about hypervisor types, virtualisation technologies.
  + What images work with what VM management environment?
* Related but not directly clear and discussed:
  + Pricing.
  + Security info.
  + Information protection.
* Consumer information:
  + End-user and brokers/meta-schedulers.
  + Also the services itself work with this information ‘locally’ or to be able to contact ‘related services’.

Standards:

* Information model context.
  + Many specifications have same core information.
  + OCCI.
  + OVF covers specification of VMs.
  + Work with DMTF.
* Information discovering.
  + OCCI has a generic discovery interface.
    - Including resource description based on key/value pairs.

Best practices:

* Are there best practices we should adopt?
  + StratusLab will address during 2nd year (federation of cloud systems).
  + Based on OpenNebula:
    - Info about the physical system.
    - Ganglia for status (monitoring).
  + From grid:
    - Information model evaluate the info captured by GLUE 2.0.
    - Difficult to upgrade information models.
      * Due to dependencies on usage/sensors.
      * Re-user experience and try.

Software maturity and availability:

* How to deliver information?
  + Traditional models like in grids now.
  + Alternative and complementary with messaging.

Priority:

* What are the priorities?
* What capabilities need to be represented?
  + Compute
    - VM appliances information.
    - Type of VMs that can be used.
  + Storage
  + Network

Work items and next steps:

* What work is needed to remove this gap?
  + Virtualisation doesn’t look like a massive big problem.
  + Issue with revising the information model.
* Analyse overlaps and differences between grid and cloud standards relating to information model discovery.

Questions:

* IB: The solution is too complex, not all the information is really needed for good scheduling. Most of the information is wrong. The lesson is to do it as simple as possible, to keep it up-to date and correct.
* Topics:
  + Physical cores are irrelevant.
  + Firewalls & Authorisation

## Key points summary

Monitoring:

* Treat VM as a black box.
  + Cannot look inside.
* Infrastructure
  + Can I start, stop, and see status of a VM through the interface.
    - Need: Nagios probes with test image (VO specific?).
    - Need: Nagios instance to host this?
* Experts level
  + Running updates on CPU load, network, i/o, etc.
  + Sensible things that can be extracted from the VMM.
    - Establish pub, sub values.

Accounting:

* Accounting for use of VM.
* VM instances can be long running & idle.
  + Need: Incremental accounting.
  + Need: To account for resources allocated.
* Need: Account for storage
  + Different life cycle than a job or running VM.
  + Physical persistent space even if used in conjunction with VM.
* Accounting for network, i/o, use,...
* Need: Infrastructure to support incremental VM and storage issues.
* Need: Understand if VMM can provide this data?
  + e.g. OpenNebula generates logs that can be used to extract information.

Management:

* Required capabilities.
  + Control the configuration of the instantiated VM.
  + Instantiate VMs as part of a cluster.
  + State discovery from as part of a cluster.
  + State discovery from the given VM handle.
  + Authentication through x.509.
  + Bulk management snapshots, exclusive use.
  + Long running VM to the source of the image.
* Need: List of capabilities in individual implementations exposed in which standards and their maturity.

Questions:

* Scheduling policies for cloud resources vs. resources allocation.

Information:

* Relating to the VMM environment.
  + Not the services running inside the image.
* Expose VMM general system state, load, and capacity.
  + Info, monitoring, measure??
* Expose the VMM trust policy for images.
* Expose enabled capabilities.
  + Allow snapshots, firewall rules, pricing, hypervisors, and image formats.
* Need: What can be delivered through the standard VMM interface and where are probes needed.
* Need: What contents can be inherited from current GLUE schema.

Other issues:

* Do we need to support the monitoring of stuff running inside the machine?
  + Establishing when a VM has started (or not).
* Standards & activities:
  + OGF accounting (storage & VM).
  + OGF: Network Mark-up of VM definition.
  + OVF+: Support.
* Billing: How to bill and who to bill?
* Policy around trust, etc.
* Common taxonomy for information, measurement and accounting.
* Validation and increasing links with user communities.
* Is there authentication on image start-up (StratusLab Marketplace)?
* Where does data reside for a VM that want to process it?
  + Difference between small volume single location and large distributed data.

Comments:

* JT: One more issue, data is more homeless in the cloud than in the grid.
* DW: It is a community specific issue; users should connect data storage services to the cloud services.
* IB: For particle physics and astrophysics, the amount of data is big, and it has to be distributed before used.

## Technical discussion: Implementation roadmap

StratusLab - Cal Loomis (CL)

* Working on many requirements listed in the summary presentations.
  + Quite reassuring.
* StratusLab is a publically accessible testbed.
  + Just send a mail for access.
  + Can be used for a testbed for whoever wants to use it.
* VM management should be in first release.
* No real issues foreseen in information systems.
* Accounting.
* Not much more is need for information systems than what we already have (GLUE, etc.)
* OVF: We are willing to work with (already in Claudia that we are using).
* Open issue:
  + How to show a VM as an infinite resource on a real machine

IGE - Helmut Heller

* Negotiating an MoU with StratusLab, people can easily instantiate Globus environment in StratusLab Cloud.
* They will make everything available on the StratusLab Marketplace.
* Venus-C and StratusLab requested this collaboration, but after that there are no limitations for other clouds.
* Plans for VM images automatic creation tools? Can’t answer to this question, too detailed.

EMI - Alberto De Meglio

* EMI works, develops, and operates in the service level, not in the VM level.
* However, some of the EMI services can be useful on the VM level too.
  + BDII as the “info service”, storing info VMMs have, so on.
  + A generic service registry, which will enable the “hooking up” of all the appliances.
* We’ll also provide, contribute, build, test grid appliances and “landscape deployments” of EMI services.
* Utilise VMs in/behind EMI services, e.g. computing elements with VM-based nodes.
* Standards: feedback on GLUE, accounting.
* Grid experience.

UK NGS - David Wallom (DM)

* UK NGS will be releasing the accounting tools (VMs and storage) through Canonical so will be available via Ubuntu.
* Currently federating clouds.
  + Happy to include StratusLab.
  + Already supporting Elixir and Clarion.
* We will run your appliances when they are ready.
* Question: Will this include summary information to see what was run?
  + DW: No, taking the amazon approach, you do not need to see my hardware.

Venus-C - Ake Edlund (AE)

* Received 60 applications for the open call.
* Committee reviewed these, and now they are 10.
* SN: What about Venus-C technical work?
  + OpenNebula, Eucalyptus, CDMI standards via EMI and NorduGrid, StratusLab.
    - SN: Is there a CDMI implementation?
    - AE: Yes, through StratusLab.

CESNET - Luděk Matyska (LM)

* Infrastructure is almost completely virtualised already (has been for a few years).
* Just offering grid interfaces.
* Putting together grid and cloud is the main difficulty.

BigGrid

* Sara has a cloud project.
  + Maybe connect that to this.
* Future resources are coming - part of that is will be partitioned for this; so we could look at that down the road.

*Discussion*

SN: How many people are using StratusLab and using OpenNebula on it’s own? The people/organisations in ( ) will be the contact points for each technology in order to investigate capabilities and functionalities.

* OpenNebula (CESGA and OpenNebula team)
* StratusLab (Cal Loomis / David O'Callaghan)
* VMware (David Wallom, OeRC)
* Azure (Ake Edlund, KTH)
* Hyper-V (STFC / CERN)
* Platform ISF (CERN)
* Eucalyptus (Ubuntu EC) (David Wallom, OeRC)
* OpenStack (CSC)
* KVM (Davide Salomoni, INFN)
* Zero writing there own

Action: EGI.eu to set up a wiki page in order to host this information and track.

In production:

* 2-3 production clouds.
* DW: Eucalyptus in production
* Spain CESGA: OpenNebula in production.

Actions:

* Nagios probes:
  + StratusLab has monitoring tools that could be integrated in Nagios probes.
  + David, Alexander, Cal to explore.
* Amazon, VMware = UK NGS?
* OpenNebula = OpenNebula team
* Explore what different capabilities are offered by different virtual manager? No one
* Accounting: VM records, storage records.
* EGI-InSPIRE has a milestone to handle this.

Who has users that are interested in use those clouds tools/services? What are the use cases already known?

Discussion around EGI vision, use cases from end-users; problem for user communities to define their own vision and provide use cases.

* IB: Would like to see a 1 page document with the vision of EGI;
* SN stated that the vision is defined in opening presentation and document, maybe needs to be extracted into the 1-page.
* JT: Explained that for the Dutch NGI works to go to users with a number of things they could do and ask them to pick up which of them they would like to be implemented or used; in front of options, users are more responsive.

Question:

* The vision is to go toward virtualised infrastructure. Does it mean that EGI can include a virtual infrastructure, or that it is based on virtual infrastructure? Is it needed?
* SN: EGI is very heterogeneous; gLite/UNICORE/ARC/Globus there will be desktop; virtualised resources are just another type.
* OS: Users just want to have a consistent user environment.

Only a few in the audience thought we should just stop with virtualisation, the majority were in favour of continuing to explore the technical implications.

Represented NGIs:

* France, Czech, Netherlands, Greece, German, UK, Poland, Sweden, Italy, Ireland, Finland.
* CZ: Will collaborate towards virtualisation; they have national program on that and there should not be too much deviation from that.
* JT/NL: Not a blank check for manpower to the project, but BigGrid will offer a part of the infrastructure for virtualisation efforts.
* SN: We have production infrastructure in place, we will NOT start to rip up the infrastructure or dissect it in anyway in order to implement any virtualisation technology.

Agreement:

* The informal agreement is that we will move forward with implementation of virtualisation, but first on a “testbed” environment.
* Question: Will EGI provide a framework for the virtual images endorsement?
* In the StratusLab MoU there was the intention to explore the sustainability of the StratusLab marketplace in the long term.

## Concluding Remarks

SN: Thanked everyone for their active participation to what proved to be very effective workshop. A short report will be produced the following weeks that will build upon the published minutes taken by the EGI.eu Policy Development Team, who was thanked as well.

A mailing list ([cloud-discuss@mailman.egi.eu](mailto:cloud-discuss@mailman.egi.eu)) will be set up in order to continue discussions during the following months leading up to the EGI Technical Forum in Lyon - Sept 2011.